

7th International Symposium on Live Subtitling and Accessibility

Universitat Autònoma de Barcelona

5-6 November 2020 (online)

PANEL 2: EMERGING AREAS: HYBRID FORMS AND INTERLINGUAL LIVE SUBTITLING.

"Real-time interlingual speech-to-text via speech recognition: mapping the field of a surging accessibility service"

Elena Davitti (University of Surrey)

Hayley Dawson (University of Surrey)

Louise Fryer (University of Surrey)

Zoe Moores (University of Surrey)

VIDEO PRESENTATION <https://youtu.be/ZLTsN4lb7nQ>

ABSTRACT

In a world where audiovisual content is produced at an unprecedented rate, there is increasing demand for live content to be made accessible in different languages for a wide audience, encompassing individuals who are deaf/hard-of-hearing and native/other-language speakers alike, and in a variety of settings, including television, conferences and other live events. Real-time interlingual speech-to-text can be used as an umbrella term to encompass a variety of technology-enabled practices to achieve this goal, including speech recognition (SR)-based methods like interlingual respeaking (IRSP).

IRSP is an emerging practice and innovative method that relies on human-machine interaction: respeakers listen to live input and simultaneously render it (with added oral punctuation) in a target language to a (speaker-dependent) SR software that turns it into written text displayed on screen. As the practice is still in its infancy, research is investigating key issues around its feasibility, the quality of the output,

the required competences, the working set-up, including configuration and location of the 'respeaking team' and ergonomics of the respeaker's workstation.

The proposed presentation will firstly characterise IRSP as a method at the crossroads of Interpreting Studies, Audiovisual Translation and Media Accessibility, adding a Human-Machine interaction dimension to it. As a practice, IRSP will be situated within the broader field of real-time interlingual speech-to-text transfer, spelling out different key dimensions (Davitti and Sandrelli 2020), including emerging set-ups and configurations. Technological development in the field of automatic speech recognition and machine translation are leading to increasingly (semi-)automated workflows to provide an interlingual real-time speech-to-text service. Different options are being tested in the quest for solutions that can streamline the process while increasing productivity and efficiency; this also raises new questions in relation to the fit-for-purposeness of the output, and the place of human input in such processes. Several possibilities will be mapped out on a continuum from 'human-centred' to 'semi-' and 'fully-automated' solutions.

After outlining some key questions emerging from the characterisation of this field and practice, based on extensive literature review as well as cross-stakeholder collaboration to identify the pressing industry needs and demands, this paper will present our recently ESRC-funded project SMART (Shaping Multilingual Access through Respeaking Technology, 2020-2022, Economic and Social Research Council UK, ES/T002530/1). This project focuses on IRSP as a human-centred form of interlingual speech-to-text transfer and aims to obtain further insights into this complex practice through experiments conducted with a population of language professionals from different backgrounds. Emphasis will be placed on the conceptual underpinning and interdisciplinary methodological approach adopted, which aims to test and correlate different variables that may be critical to IRSP performance. Findings will ultimately inform training to upskill language professionals with a view to adding IRSP to their profile.